



Grupo de investigación Ecología de Zonas Áridas

CENTRO ANDALUZ PARA LA EVALUACIÓN Y SEGUIMIENTO DEL CAMBIO GLOBAL

Modelling Workshops

GBIF

¿Qué modelo elegir?

- **BIOCLIM.**
- **DOMAIN**
- **GARP**
- **MaxEnt**

Elisa Liras
Dpto. Biología Vegetal y Ecología
Universidad de Almería
eliras@ual.es

"Para la modelización del nicho ecológico necesitamos herramientas informáticas y estadísticas, que utilizando registros de las especies georreferenciados y variables ambientales, produzcan representaciones formales del espacio ecológico (nicho) sobre el espacio geográfico (distribución)".

8. ¿Qué modelo elegir?

Hay varios algoritmos diferentes para estimar la probabilidad de presencia de una especie (y la ausencia) en función de un conjunto de variables predictoras.

La clave está en identificar posibles relaciones no lineales en un ambiente multidimensional.

En este taller vamos a ver las diferencias entre:

- **BIOCLIM**
- **DOMAIN**
- **GARP**
- **MaxEnt**

No obstante, hay muchos otros métodos, como GLMs, GAM, Redes Neuronales,…

8. ¿Qué modelo elegir?

Modelos de nicho ecológico

Método	Tipo de modelo	Datos[1]	Software
BIOCLIM	modelo de envuelta	p	DIVA-GIS
BRT	árboles de decisión amplificados	pa	R, gbm package
BRUTO	regresión, implementación rapida de gam	pa	R & Splus, mda package
DKP-GARP	reglas para algoritmos genéticos, versión de escritorio	pa	DesktopGarp
DOMAIN	distancia multivariante	p	DIVA-GIS
GAM	regresión: modelos aditivos de generalización	pa	S-Plus, GRASP
GDM	modelos de disimilaridad generales, utiliza datos de comunidad	pacomm	programa especializado que utiliza ArcView y Splus
GDM-SS	modelos de disimilaridad generales, implementación para especies	pa	igual que GDM
GLM	regresión, modelos lineales generalizados	pa	S-Plus, GRASP
LIVES	distancia multivariante	p	programa especializado
MARS	regresión, regresiones multivariantes adaptativas	pa	R, mda package
MARS-COMM	implementación de MARS para datos de comunidad	pacomm	como MARS
MARS-INT	implementación de MARS para permitir interacciones	pa	como MARS
MAXENT	máxima entropía	pa	Maxent
MAXENT-T	máxima entropía con valores de frontera	pa	Maxent
OM-GARP	versión de código abierto de GARP	pa	nueva versión de GARP

[1] *p = solo presencia; pa = datos de presencia y ausencia; comm = datos de comunidad.*

(basado en Elith et al. 2006)

8. ¿Qué modelo elegir?

En general, hay 2 aspectos importantes que pueden condicionar el algoritmo a utilizar

1. **Tipos de datos que queremos analizar** (presencias / ausencias)

2. **Tipo de salida que queremos**… qué a su vez depende de la aplicación que le queremos dar al resultado (distribución / nicho ecológico)

1. Tipos de datos que queremos analizar (presencias / ausencias)

Modelos de nicho ecológico

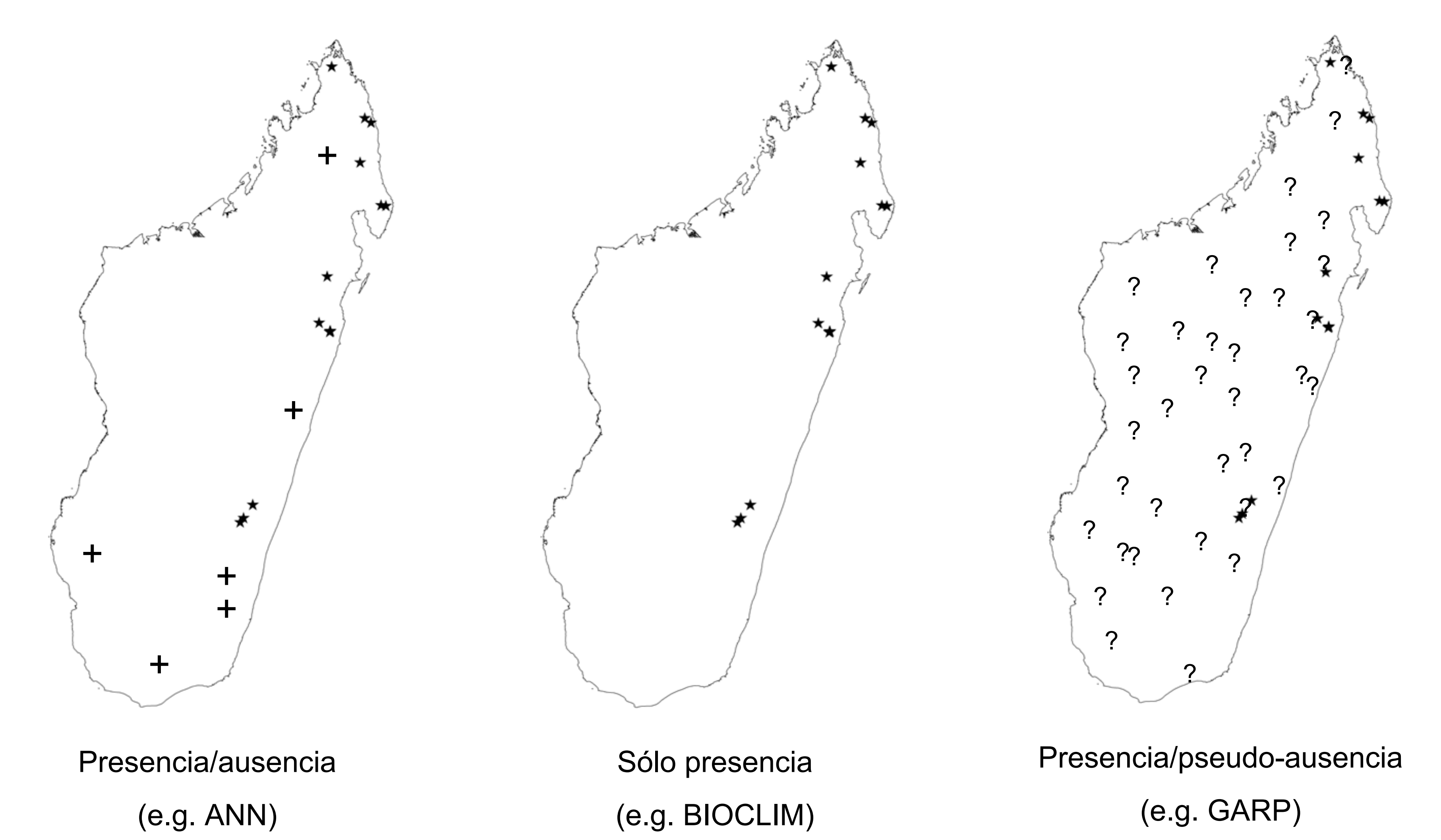
Método	Tipo de modelo	Datos[1]	Software
BIOCLIM	modelo de envuelta	p	DIVA-GIS
BRT	árboles de decisión amplificados	pa	R, gbm package
BRUTO	regresión, implementación rapida de gam	pa	R & Splus, mda package
DKP-GARP	reglas para algoritmos genéticos, versión de escritorio	pa	DesktopGarp
DOMAIN	distancia multivariante	p	DIVA-GIS
GAM	regresión: modelos aditivos de generalización	pa	S-Plus, GRASP
GDM	modelos de disimilaridad generales, utiliza datos de comunidad	pacomm	programa especializado que utiliza ArcView y Splus
GDM-SS	modelos de disimilaridad generales, implementación para especies	pa	igual que GDM
GLM	regresión, modelos lineales generalizados	pa	S-Plus, GRASP
LIVES	distancia multivariante	p	programa especializado
MARS	regresión, regresiones multivariantes adaptativas	pa	R, mda package
MARS-COMM	implementación de MARS para datos de comunidad	pacomm	como MARS
MARS-INT	implementación de MARS para permitir interacciones	pa	como MARS
MAXENT	máxima entropía	pa	Maxent
MAXENT-T	máxima entropía con valores de frontera	pa	Maxent
OM-GARP	versión de código abierto de GARP	pa	nueva versión de GARP

[1] *p = solo presencia; pa = datos de presencia y ausencia; comm = datos de comunidad.*

(basado en Elith et al. 2006)

1. Tipos de datos que queremos analizar (presencias / ausencias)

Modo de empleo de los datos de las especies



8. ¿Qué modelo elegir?

1. Tipos de datos que queremos analizar (presencias / ausencias)

¿Qué es una ausencia?

Una localidad puede ser clasificada como ausencia por:

1. La especie no fue detectada, pero se encuentra allí



2. La especie está ausente, pero el ambiente es propicio. No hay ausencia de su nicho

'Falsa ausencia'

3. La especie está ausente y el ambiente es inapropiado



Presentación original: Martínez-Meyer

1. Tipos de datos que queremos analizar (presencias / ausencias)

Además hay que recordar que…

El resultado depende más de la calidad de los datos biológicos de entrada que del algoritmo utilizado.

MORE COMPLEX DISTRIBUTION MODELS OR MORE REPRESENTATIVE DATA?

JORGE M. LOBO

Dpto. de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), c/José Gutiérrez Abascal, 2 – 28006, Madrid, Spain. e-mail: mcnj117@mncn.csic.es

Abstract.— Distribution models for species are increasingly used to summarize species' geography in conservation analyses. These models use increasingly sophisticated modeling techniques, but often lack detailed examination of the quality of the biological occurrence data on which they are based. I analyze the results of the best comparative study of the performance of different modeling techniques, which used pseudo-absence data selected at random. I provide an example of variation in model accuracy depending on the type of absence information used, showing that good model predictions depend most critically on better biological data.

Key words.—distribution models, model reliability, pseudo-absences, conservation usefulness.

Biodiversity Informatics, 5, 2008, pp.14-19

2. Tipo de salida que queremos (mapa binario / probabilístico)

Depende de la aplicación que le queramos dar al modelo

1. **Salida binaria:** no nos da información sobre la relación de la especie con las variables ambientales ni de las variables ambientales entre sí (caja negra)

   → ***funcionan muy bien, y nos sirven cuando sólo estamos interesados en la distribución potencial de la especie (ej. Riesgo de invasión biológica, búsqueda de nuevas localidades de presencia)***

2. **Salida probabilística:** informa sobre la relación de la especie con las variables ambientales y la relación entre variables

   → ***nos dan información sobre las variables que determinan el nicho ecológico de la especie (cuáles son más importante). Esta salida es necesaria cuando lo que buscamos es comprender la ecología de la especies.***

2. Tipo de salida que queremos (mapa binario / probabilístico)

Modelos de nicho ecológico

Método	Tipo de modelo	Datos[1]	Software
BIOCLIM	modelo de envuelta	p	DIVA-GIS
BRT	árboles de decisión amplificados	pa	R, gbm package
BRUTO	regresión, implementación rapida de gam	pa	R & Splus, mda package
DKP-GARP	reglas para algoritmos genéticos, versión de escritorio	pa	DesktopGarp
DOMAIN	distancia multivariante	p	DIVA-GIS
GAM	regresión: modelos aditivos de generalización	pa	S-Plus, GRASP
GDM	modelos de disimilaridad generales, utiliza datos de comunidad	pacomm	programa especializado que utiliza ArcView y Splus
GDM-SS	modelos de disimilaridad generales, implementación para especies	pa	igual que GDM
GLM	regresión, modelos lineales generalizados	pa	S-Plus, GRASP
LIVES	distancia multivariante	p	programa especializado
MARS	regresión, regresiones multivariantes adaptativas	pa	R, mda package
MARS-COMM	implementación de MARS para datos de comunidad	pacomm	como MARS
MARS-INT	implementación de MARS para permitir interacciones	pa	como MARS
MAXENT	máxima entropía	pa	Maxent
MAXENT-T	máxima entropía con valores de frontera	pa	Maxent
OM-GARP	versión de código abierto de GARP	pa	nueva versión de GARP

[1] *p = solo presencia; pa = datos de presencia y ausencia; comm = datos de comunidad.*

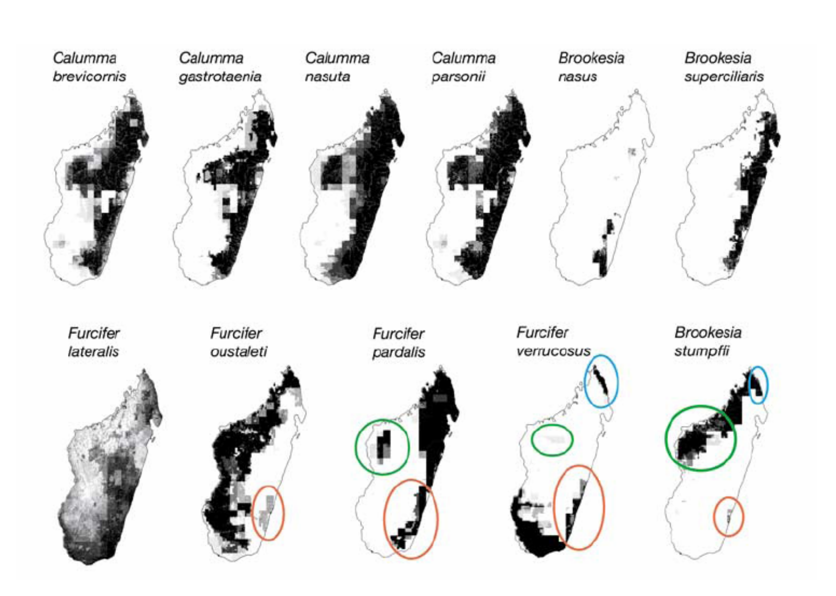
(basado en Elith et al. 2006)

2. Tipo de salida que queremos (mapa binario / probabilístico)

2 ejemplos anteriores…

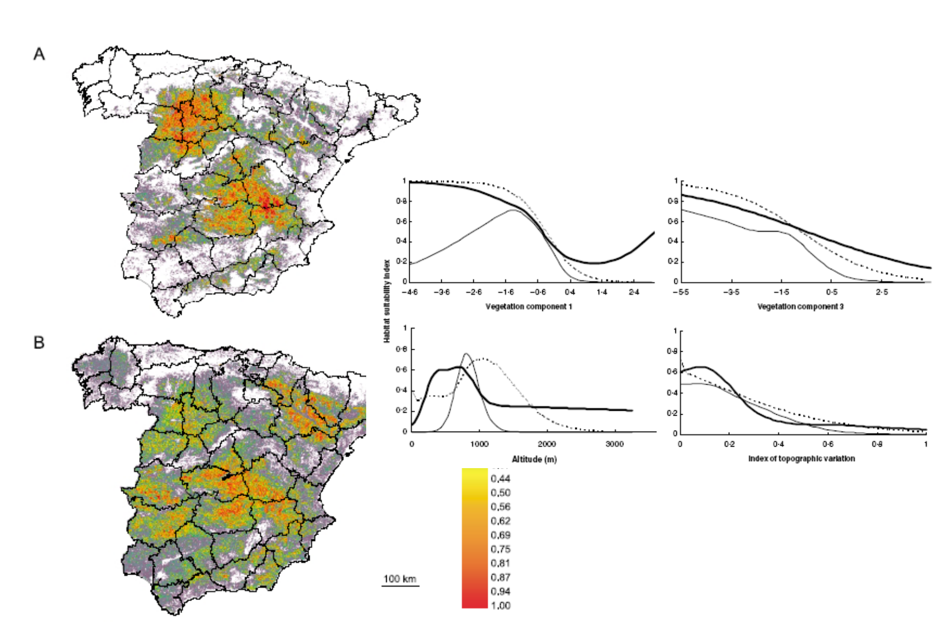
Camaleones en Madagascar:

nos importa la predicción geográfica (buscan nuevas localidades de presencia)



Aves esteparias en la Península Ibérica:

nos importan las curvas de respuesta para las variables (buscan entender la ecología de la especie)



Entonces…

Veamos que hace cada uno de los algoritmos de modelización con los que vamos a trabajar en el taller

8. ¿Qué modelo elegir?

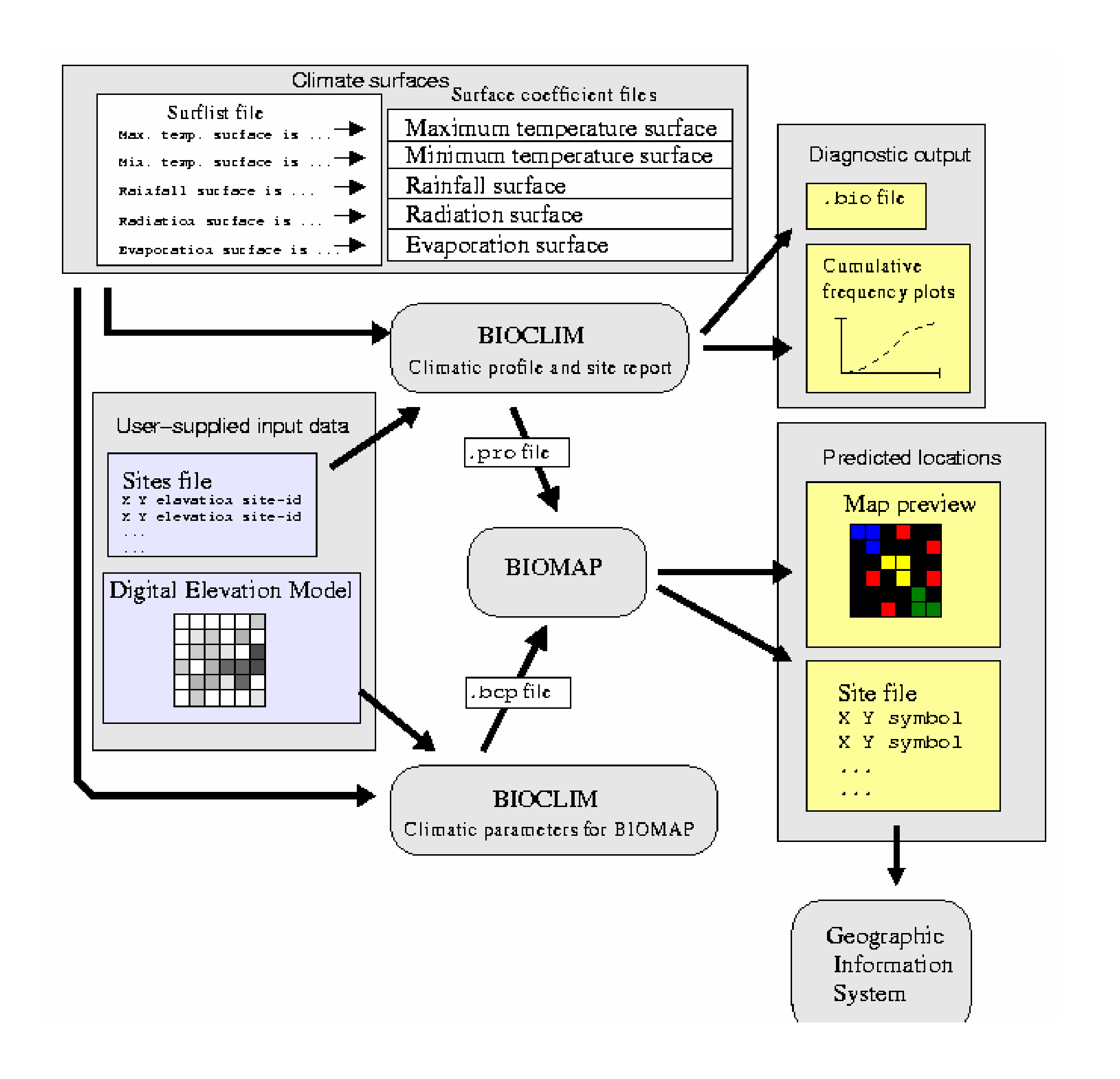
Modelos de nicho ecológico

Método	Tipo de modelo	Datos[1]	Software
BIOCLIM	modelo de envuelta	p	DIVA-GIS
BRT	árboles de decisión amplificados	pa	R, gbm package
BRUTO	regresión, implementación rapida de gam	pa	R & Splus, mda package
DKP-GARP	reglas para algoritmos genéticos, versión de escritorio	pa	DesktopGarp
DOMAIN	distancia multivariante	p	DIVA-GIS
GAM	regresión: modelos aditivos de generalización	pa	S-Plus, GRASP
GDM	modelos de disimilaridad generales, utiliza datos de comunidad	pacomm	programa especializado que utiliza ArcView y Splus
GDM-SS	modelos de disimilaridad generales, implementación para especies	pa	igual que GDM
GLM	regresión, modelos lineales generalizados	pa	S-Plus, GRASP
LIVES	distancia multivariante	p	programa especializado
MARS	regresión, regresiones multivariantes adaptativas	pa	R, mda package
MARS-COMM	implementación de MARS para datos de comunidad	pacomm	como MARS
MARS-INT	implementación de MARS para permitir interacciones	pa	como MARS
MAXENT	máxima entropía	pa	Maxent
MAXENT-T	máxima entropía con valores de frontera	pa	Maxent
OM-GARP	versión de código abierto de GARP	pa	nueva versión de GARP

[1] *p = solo presencia; pa = datos de presencia y ausencia; comm = datos de comunidad.*

(basado en Elith et al. 2006)

BIOCLIM



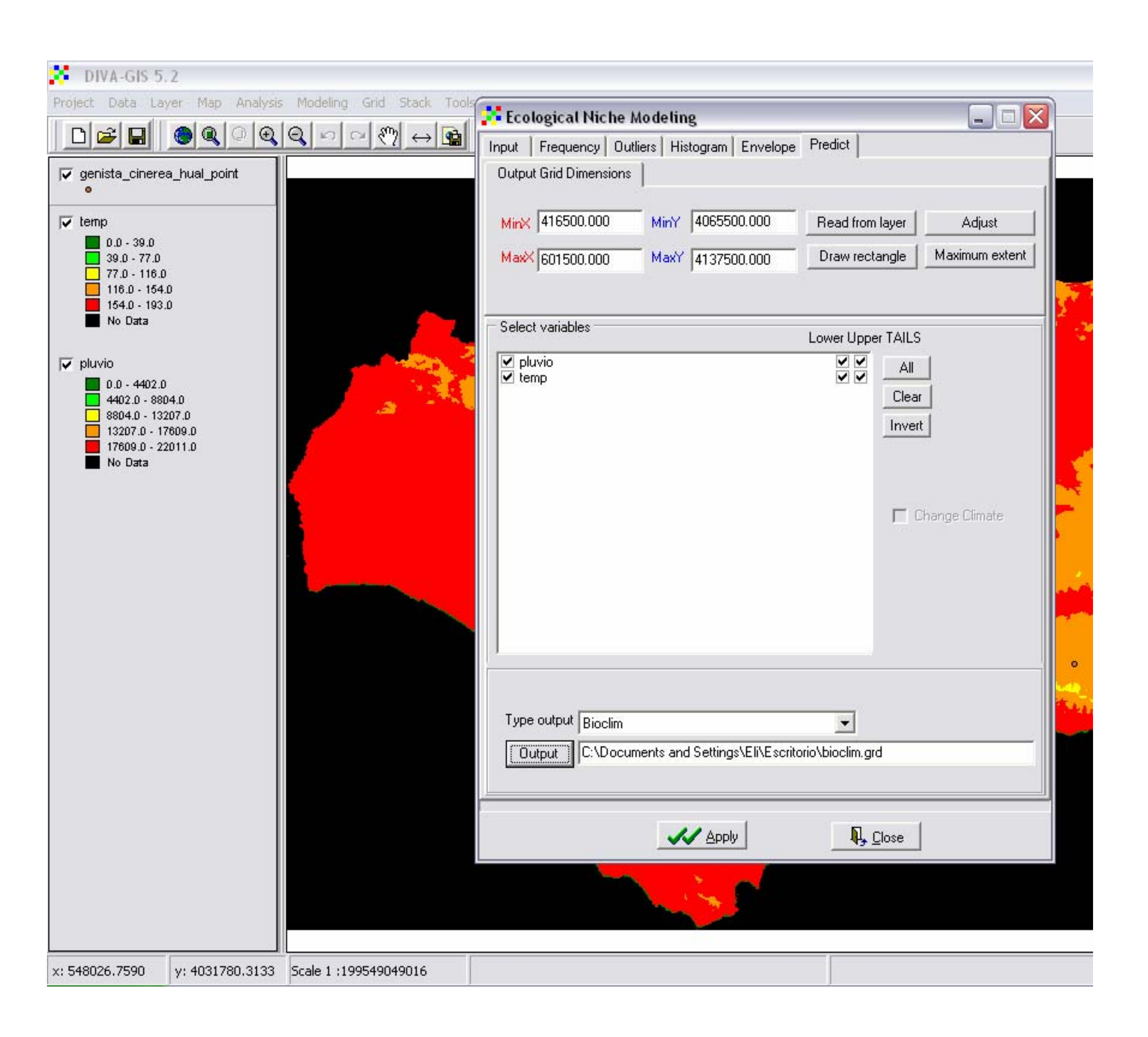
Algoritmo:
Envoltura bioclimática

Datos de entrada:
Sólo presencia

Productos:
Mapas binarios

Desventajas:
No tiene en cuenta las posibles relaciones entre las variables utilizadas.

DOMAIN



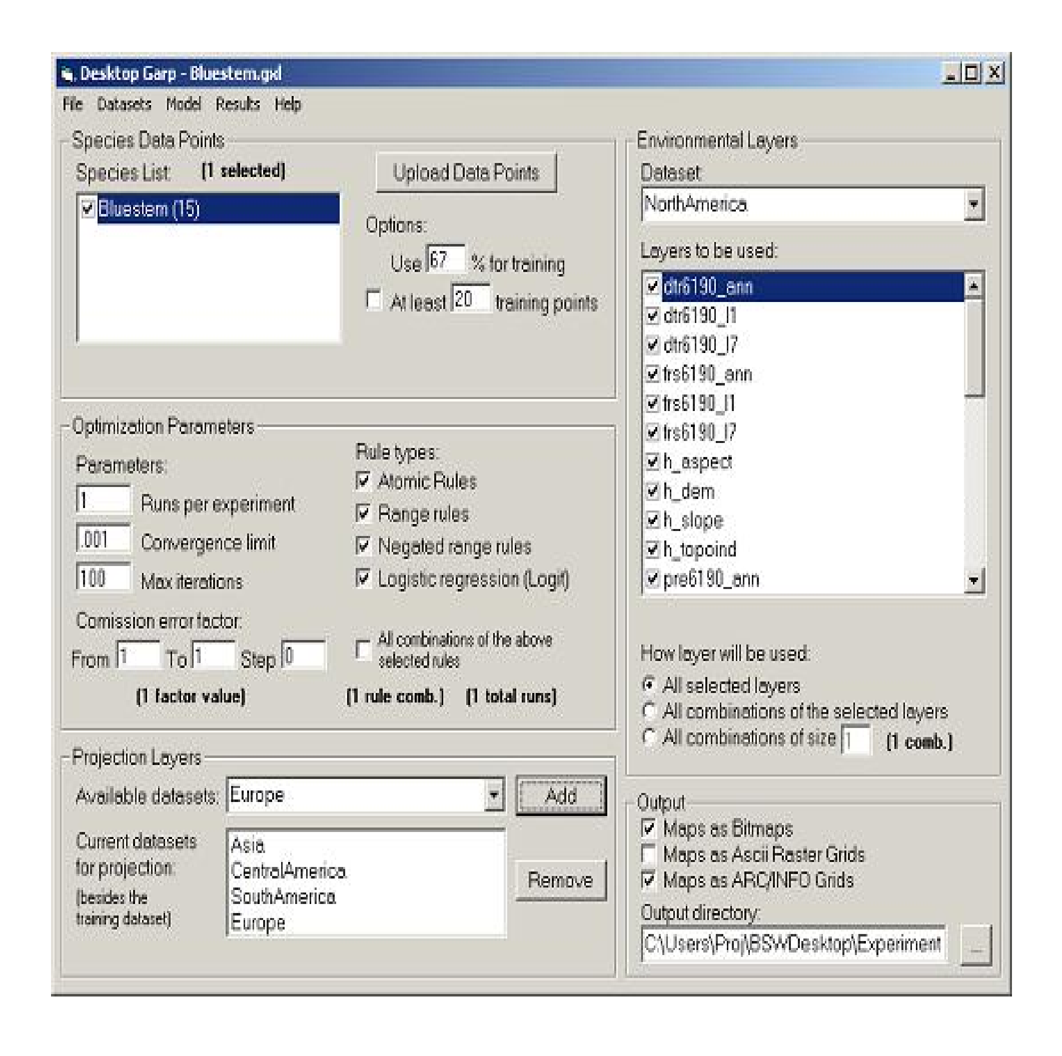
Algoritmo:
Parámetro de Gower (similitud)

Datos de entrada:
Sólo presencia

Productos:
Mapas probabilísticos

Desventajas:
No tiene en cuenta las posibles relaciones entre las variables utilizadas

GARP



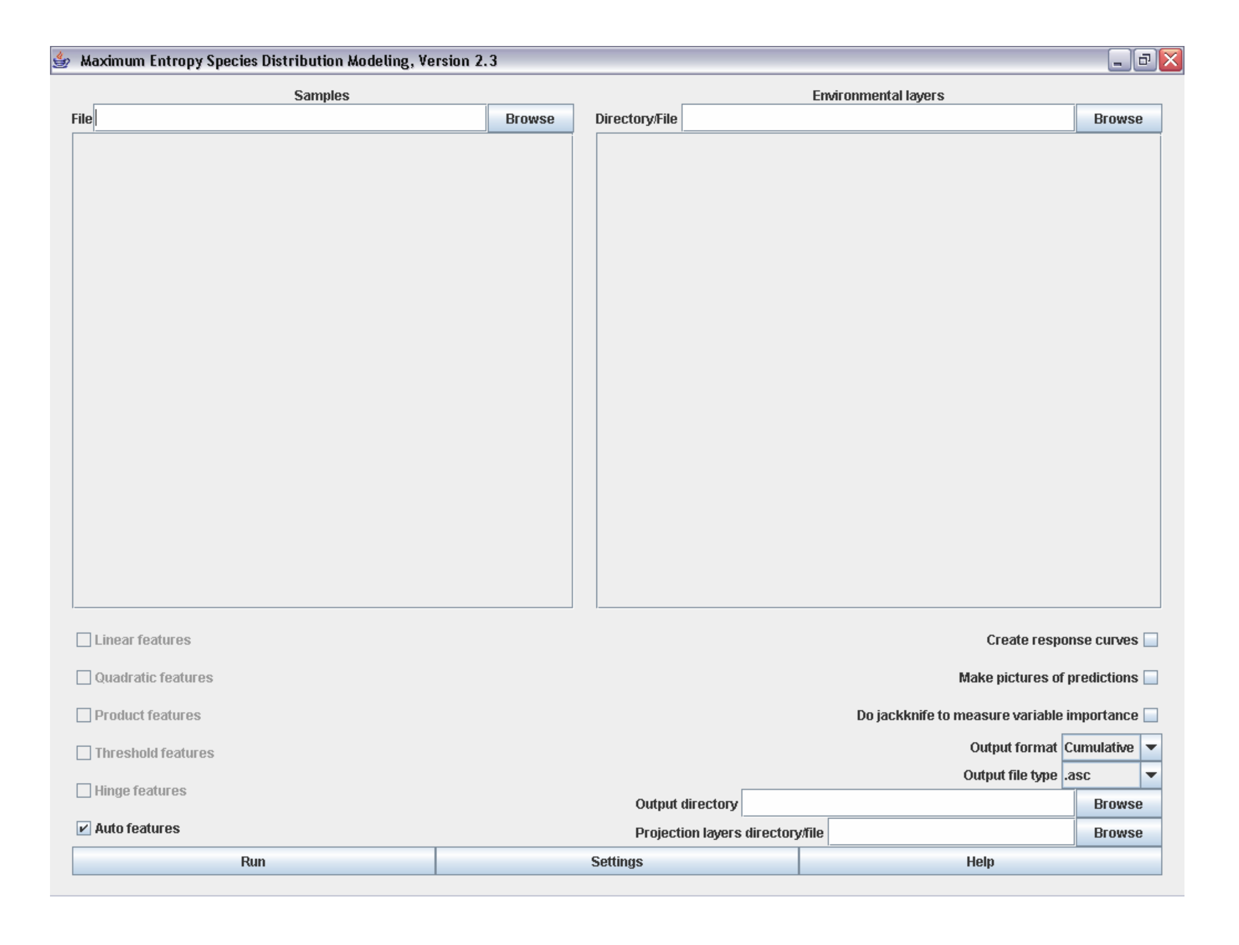
Algoritmo:
Algoritmo genético

Datos de entrada:
Sólo presencia

Productos:
Mapas binarios

Desventajas:
Actúa como una caja negra… no da curvas de respuesta

MaxEnt



Algoritmo:
Distribución de Máxima Entropía

Datos de entrada:
Sólo presencia

Productos:
Mapas probabilísticos
Variables explicativas
Curvas ROC

Desventajas:
Tiende a sobreajustar el modelo

Un comentario final

El modelo de nicho ecológico WhyWhere tiene muchas críticas… al márgen de la estadística, es importante tenerlo en cuenta a la hora de enviar un artículo a publicar

ECOLOGICAL MODELLING 203 (2007) 527–530

ELSEVIER

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel

ECOLOGICAL MODELLING

Short communication

Why not WhyWhere: The need for more complex models of simpler environmental spaces

A. Townsend Peterson

Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, Kansas 66045, USA

ARTICLE INFO

Article history:
Received 24 April 2006
Received in revised form

ABSTRACT

WhyWhere has recently been introduced by Stockwell as a new tool for ecological niche modeling or species distribution modeling. I address two features of WhyWhere and its presentation: (1) the assertion that hundreds of environmental data layers are necessary to

Entoces…¿Qué modelo elegir?

Actualmente, existen del orden de 12 a 15 modelos de nicho ecológico (e.g. Elith et al. 2006,), que se diferencian entre sí en la forma en que, matemáticamente, estiman el nicho potencial de las especies modeladas.

En general, no existe uno mejor o peor, sino que la clave está en entender qué hace cada uno de los software y determinar cual es más conveniente para nuestra aplicación.

Esta gran “inflación” de métodos matemáticos para modelizar el nicho de las especies ha hecho que empiecen a surgir nuevos procedimientos matemáticos que integren de alguna manera los resultados obtenidos por los diferentes algoritmos.

En este sentido destacamos la denominada “predicción conjunta” (ensemble forecastig) (Araujo & New 2007).

Liras et al., 2008

ELSEVIER

Review

TRENDS in Ecology and Evolution Vol.22 No.1

Full text provided by www.sciencedirect.com



Ensemble forecasting of species distributions

Miguel B. Araújo[1] and Mark New[2]

[1] Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, CSIC, C/Gutiérrez Abascal, 2, 28006, Madrid, Spain
[2] Climate Research Laboratory, Oxford University Centre for the Environment, South Parks Road, Oxford, UK, OX1 3QY

Concern over implications of climate change for biodiversity has led to the use of bioclimatic models to forecast the range shifts of species under future climate-change scenarios. Recent studies have demonstrated that projections by alternative models can be so variable as to compromise their usefulness for guiding policy decisions. Here, we advocate the use of multiple models within an ensemble forecasting framework and describe alternative approaches to the analysis of bioclimatic ensembles, including bounding box, consensus

in studies comparing alternative techniques to assess potential climate change-induced shifts in the distributions of European plants [15], amphibians and reptiles [16], and British breeding birds [17]. These results challenge the common practice of relying on one single method to make forecasts of the responses of species to climate change scenarios or, if one accepted a more sceptical view, the usefulness of bioclimatic modelling in general for climate change impact studies.

Such variability in forecasts is not surprising given that

Muchas gracias